DOCUMENT RESUME

ED 354 259 TM 019 519

AUTHOR Schenck, E. Allen
TITLE Summary of Recommendations Regarding Chapter 1 Assessment from Selected Sources. Prepared for the Advisory Committee on Testing in Chapter 1.
INSTITUTION Chapter 1 Technical Assistance Center, Denver, CO. Region E.
SPONS AGENCY Department of Education, Washington, DC.
REPORT NO TAC-B-292
PUB DATE 28 Sep 90
NOTE 14p.
PUB TYPE Information Analyses (070) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Accountability; *Compensatory Education; *Economically Disadvantaged; *Educational Assessment; Educationally Disadvantaged; Educational Policy; Educational Technology; Elementary Secondary Education; Federal Legislation; *Federal Programs; *Guidelines; Literature Reviews; *Policy Formation; Program Effectiveness; Test Use
IDENTIFIERS *Education Consolidation Improvement Act Chapter 1; Hawkins Stafford Act 1988; Performance Based Evaluation; Testimony

ABSTRACT

Six documents were reviewed and recommendations relevant to the work of the Advisory Committee on Testing in Chapter 1 were summarized. A brief description of a framework for discussing different uses of assessment is followed by specific recommendations from each of the following sources: (1) "Working Paper: Reauthorization of the Hawkins-Stafford Amendments of 1988" (Council of Chief State School Officers); (2) "Summary of Public Comments Regarding the Reauthorization of Elementary and Secondary Education Programs" (Pelavin Associates, Inc.); (3) "Chapter 1 Reauthorization: Testimony before the Advisory Committee on Testing in Chapter 1" (M. H. Kean); (4) "Chapter 1 Assessment Issues" (W. L. Padia); (5) "Testing in American Schools: Asking the Right Questions" (U.S. Congress, Office of Technology Assessment); and (6) "Regional Hearing Summaries" (U.S. Department of Education, Office of Management and Budget). Nineteen recommendations are made in the areas of test use, accountability, and program effectiveness. Attention is given to testing practices, performance measures, and new technology. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TAC B-292

ED354259

SUMMARY OF RECOMMENDATIONS REGARDING CHAPTER 1 ASSESSMENT FROM SELECTED SOURCES

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: TM
In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view. UD

Prepared for the

Advisory Committee on Testing in Chapter 1

by

E. Allen Schenck
RMC Research Corporation

September 28, 1990
Washington, DC

BEST COPY AVAILABLE

METHODOLOGY

The author was asked to review the following six documents and to summarize any explicit or implicit recommendations in them which were relevant to the charge of the Advisory Committee on Testing in Chapter 1:

> Council of Chief State School Officers, Hawkins-Stafford Reauthorization Task Force, *Working Paper: Reauthorization of the Hawkins-Stafford Amendments of 1988*, September 1, 1992. **[CCSSO]**
>
> Deich, S.G., Sherman, J.D., Amstutz, W. and Shiffman, J., *Summary of Public Comments Regarding the Reauthorization of Elementary and Secondary Education Programs*, Pelavin Associates, Inc., April 30, 1992. **[SPC]**
>
> Kean, M.H., *ESEA Chapter 1 Reauthorization: Testimony Before the Advisory Committee on Testing in Chapter 1*, Association of American Publishers, August 12, 1992. **[Kean]**
>
> Padia, W.L., *Chapter 1 Assessment Issues*, California Department of Education, August 12, 1992. **[Padia]**
>
> U.S. Congress, Office of Technology Assessment, *Testing in American Schools: Asking the Right Questions*, OTA-SET-519 (Washington, DC: U.S. Government Printing Office, February 1992). **[OTA]**
>
> U.S. Department of Education, Office of Management and Budget, *Regional Hearing Summaries*, April 27, 1992. **[RHS:City]**

The results of this review are presented in three parts below. First there is a brief description of a framework for discussing different uses of assessment. Second, specific recommendations from each of the above six sources are presented. Each recommendation is linked to its source by a code enclosed in square brackets, e.g., [OTA-89]. The code for each source can be found at the end of the citations listed above. The number attached to each code represents the page of the document from which the recommendation was taken, e.g., [OTA-89] indicates page 89 from the Office of Technology report.

The selection of recommendations was based on the author's judgement. These recommendations should not be taken as direct quotes, although most are very close to the original language. The authors of the original documents might not agree with the selection or wording of what is presented here. An attempt was made to organize the recommendations into a useful set of categories. Placement in some cases was arbitrary. Occasionally, a single recommendation may appear in more than one category.

Following the specific recommendations is a list of summary recommendations created by the author based on what the specific recommendations seem to be saying. Others might not interpret the specific recommendations in the same way. These summary recommendations are numbered only to facilitate discussion.

FRAMEWORK FOR DISCUSSING USES OF ASSESSMENT

GENERAL USES OF ASSESSMENT

The uses of educational tests can be grouped into three areas: [OTA-10]

- to aid teachers and students in the conduct of classroom learning
- to monitor systemwide educational outcomes
- to inform decisions about student selection, placement and credentialing

See Table 6-3. [OTA-195]

Table 6-3—Functions of Tests: What Designs Are Needed?

	Classroom instructional guidance	System monitoring	Selection, placement, and certification
Who needs to be described	Individuals	Groups of students	Individuals
"Stakes" or consequences attached	Low	High or low	High
Characteristics of the test needed			
Comparability of information	Low	High	High
Impartial scoring (not teachers)	No	Yes	Yes
Standardized administration	No	Yes	Yes
Type of information needed			
Detailed v. general	Detailed	General	General
Frequency	Frequently during a single school year	Once a year or less	Once a year or less
Results needed quickly	Yes	No	No
Technical requirements			
Need for high test reliability (internal consistency and stability)	Can vary	Depends on size of group	Very high
Type of validity evidence	Content	If low stakes: content If high stakes: content and construct	Content Additional validity evidence must be demonstrated for the specific purpose (e.g., certification = criterion validity, selection = predictive validity)

SOURCE: Office of Technology Assessment, 1992; adapted from Lauren B. Resnick and Daniel P. Resnick, "Assessing the Thinking Curriculum: New Tools for Educational Reform," paper prepared for the National Commission on Testing and Public Policy, August 1989. (To appear in B.R. Gifford and M.C. Connor (eds.), *Future Assessments: Changing Views of Aptitude, Achievement, and Instruction* (Boston, MA: Kluwer Academic Publishers, in press).)

Author's Alignment of Chapter 1 Uses with OTA's Framework
System Monitoring • national accountability (TIERS) • state evaluation (biennial report) • local program evaluation • progress in the regular program • schoolwide project accountability • determine need for school program improvement plan • determine levels of need for Chapter 1 services in schools, grades and subjects Selection and Placement • identify and select eligible students • identify students not progressing adequately Instructional Guidance • diagnosis of individual student needs

SPECIFIC RECOMMENDATIONS REGARDING USES OF ASSESSMENT

GENERAL CHAPTER 1 USE

Allow SEAs and LEAs to use a variety of assessment methods that reflect their own goals for Chapter 1 in order to modify programs, place students, target schools for program improvement, decide on continuation of schoolwide projects, etc. [OTA-34]

Reduce the reliance of Chapter 1 assessment and evaluation on NRTs. [OTA-90]

Reduce the amount of testing required. [OTA-90]

Decrease emphasis on NRTs by scaling down the roles testing plays in Chapter 1 programs, reducing the number of areas and students tested, and reducing the frequency of testing. [RHS:LA-2]

Give states the flexibility to tailor their assessment strategy to the requirements for selection, needs assessment, identification of students, and evaluation. Each state's assessment plan could be negotiated with USDE. [Padia-5]

CHAPTER 1 ACCOUNTABILITY/EVALUATION

GENERAL

National and state/local Chapter 1 program administrators' data needs are different and not necessarily well met by NRTs. [OTA-89]

For accountability, multiple assessments should be used to determine student achievement and program effectiveness. [CCSSO-18]

There are existing state assessment systems that should be taken advantage of, and state should be charged to develop better systems for accountability. SEAs should be required to have an assessment system which includes a mix of assessments from which to make decisions about students and programs. [CCSSO-19]

NRTs should not be used for accountability or selection. [RHS:DC-3]

Should assessment of Chapter 1 students be required annually and in each grade? (7 yes - 4 no) [SPC-50]

- Yes - for determining special needs.
- Yes - give quarterly.
- Fall to spring is better.
- Yes - if have to have gain scores.

Should student assessment criteria other than standardized tests be used? (10 yes - 0 no) [SPC-51]

- Would like to see improvement in the classroom as the main measure of success.

Demonstrate student performance by supplementing NRTs with alternatives like writing assessment. [Kean-16]

Consider technical as well as policy implications of comparing or equating different assessment instruments even if all the instruments are aligned to the same content standard. [Kean-18]

Do not require matched pretest and posttest scores. High mobility creates potential for unrepresentative results. [Padia-6]

National Accountability

Conduct separate national assessment of Chapter 1 which would allow sampling, less frequent testing for students, less time spent testing by teachers and other school personnel, and more control over data quality. [OTA-35]

The national component of a state's assessment system would include standardized criteria or NRTs using a statewide sample (or as part of a national sample) for program accountability. [CCSSO-20]

Test a sample of Chapter 1 students. [RHS:LA-2]

Administer tests less frequently at specific grades with measures of performance. [RHS:Seattle-3]

Should testing be aligned with national assessments measuring performance less frequently and at specific grade levels? (4 yes - 2 no) [SPC-51]

- Yes - use NRT data at grades 4, 8, and 12.
- No - national assessments are not uniform, they are voluntary.

Allow statewide sampling of Chapter 1 programs and student outcomes in order to provide national comparisons. [SPC-52]

Obtain data for a national evaluation through a national sampling at selected grade levels, perhaps through NAEP or another national testing system. [Padia-7]

Tests used in a national evaluation should incorporate authentic, performance-based measures and should yield the percentage of students achieving specified performance standards. [Padia-7]

Evaluating State/Local Program Effectiveness

Consider USDE funding of research and development on better assessment practices at the state and local level. [OTA-34]

The identification of schools for program improvement and the evaluation of schoolwide projects has required more of NRTs than they can, perhaps, provide while also attaching greater consequences to their results. [OTA-89]

The state component of a state's assessment system would include district or state developed CRTs (including performance-based assessments) based on state outcomes and standards of performance. [CCSSO-20]

The local component of a state's assessment system would include other indicators of performance which LEAs and schools could use to measure individual student achievement such as portfolios, observed performance and participation rates. [CCSSO-20]

Require states to develop their own assessment techniques with stringent outcome measures and performance standards that could be aligned with state and local assessment practices. [RHS:DC-3]

Work with SEAs and LEAs to develop viable alternatives to NRTs to determine outcomes of Chapter 1 programs. [SPC-52]

Use multiple measures which reflect the breadth and depth of desired outcomes. [Kean-15]

SEAs and Committees of Practitioners should be responsible for setting state standards of performance for Chapter 1 programs at selected grade levels, with attainment measured by state or nationally developed tests. Permit and encourage tests with strong performance assessment components. Encourage LEAs to set standards for other grades. Allow NRTs until high-quality performance measures are available. Encourage the use of multiple indicators. [Padia-9]

Determining Need for School Program Improvement Plan

The identification of schools for program improvement and the evaluation of schoolwide projects has required more of NRTs than they can, perhaps, provide while also attaching greater consequences to their results. [OTA-89]

A better match is needed between the goal of improving the quality of local Chapter 1 programs and the tools used to measure progress toward that goal. [OTA-89]

Modify Chapter 1 program improvement requirements to allow other desired outcomes to determine buildings required to implement program improvement plans. [SPC-52]

Should the use of a particular measure of program effectiveness be required for identifying schools for program improvement? (4 yes - 6 no) [SPC-59]

- Writing samples, portfolios, basal unit tests and classroom teacher surveys should be allowed for measuring substantial progress toward desired outcomes.
- Multiple indicators (NRTs and desired outcomes) should be used for targeting program improvement sites, using the preponderance of evidence approach.
- Allow use of other desired outcomes in conjunction with aggregate performance.
- LEAs should have the flexibility to choose between standardized testing and desired outcomes.

Should student assessment criteria, other than standardized tests, be used for program improvement? (9 yes - 0 no) [SPC-64]

- Teacher evaluation should be used.
- School districts should design their own assessment plans.
- There should be a greater emphasis on desired outcome measures.
- Performance assessment should be used.

A local committee should set goals, standards, and measures for determining the success of the Chapter 1 program. Results should be reported to the LEA and SEA. SEA should provide technical assistance in skills of setting high and reasonable goals and standards and selecting appropriate measures and instructional practices. SEA should provide incentives for LEAs taking on high goals and innovative practices. [SPC-65]

A broad view of assessment is needed for gauging improvement. Training and monitoring are important in setting performance standards. [SPC-66]

Consider most appropriate ways to measure progress and determine the need for program improvement, such as supplementing NCE gains with other measures and judging Chapter 1 students relative to other Chapter 1 students instead of all students. [Kean-18]

Minimize disincentives for setting high standards and teaching higher order skills. [Kean-18]

Determining Levels of Need for Services

Funding should be stable for five years instead of a function of test scores. [RHS:Seattle-3]

Other Chapter 1 Uses of Assessment

Identification/Selection

Tests used for selection or placement should be designed and validated for that purpose. [OTA-185]

Tests designed to be used as feedback mechanisms to inform the learning process [referring to NRTs] should not be used to make significant decisions about an individual's educational career unless additional evidence can be provided substantiating this use. [OTA-185]

Let the teachers decide who is in the Chapter 1 program. Use portfolios to determine students' abilities. [RHS:DC-3]

Methods of identifying Chapter 1 students should be modified - alternatives to NRTs should be developed and utilized. [SPC-14]

De-emphasize NRTs for pre-Kindergarten through 3rd grade as a means for selecting students for Chapter 1. Emphasize the use of developmental assessments, readiness tests, portfolios, checklists, and documented teacher observations. [Padia-8]

Instructional Monitoring

One of the dangers in relaxing technical standards for classroom tests is that the use of the scores cannot be restricted or monitored appropriately once they are obtained. [OTA-196]

Provide for periodic monitoring of student progress. [Kean-15]

General Assessment Issues

Relationship of Assessment to Curriculum

Stimulate the development of assessment methods more suitable to the goals of Chapter 1. [OTA-90]

Avoid requirement to measure achievement with NRTs which result in schools delivering a low-level skills-based curriculum to disadvantaged students. [RHS:LA-2]

NRTs measure isolated versus complex skills. [RHS:Seattle-3]

Should Chapter 1 assessment practices be aligned with current movement to develop national standards for all children? (yes/no count missing) [SPC-48]

- Should focus on advanced skills and complex tasks.
- Develop guidelines for portfolio assessment to supplement NRTs.
- Should be aligned with local assessment.
- Allow for change as state and national standards are pu in place.
- Identify Chapter 1 students in NAEP and measure gap.
- Emphasize other desired outcomes as much or more than NRTs.
- Allow states to set standards for other desired outcomes.

Develop and expand training and technical assistance to LEA teachers and administrators to improve the link between instruction and assessment. [SPC-52]

Articulate assessment information to national education goals, national curriculum standards, and to state and local assessment approaches. [Kean-15]

Continue to assess both basic and higher order skills. [Kean-16]

Good Testing Practices

Emphasize adherence to good testing practices, including validity, reliability, fairness, security, especially for uses outside the classroom. [Kean-15]

Use different tests for different purposes. [Kean-15]

Present assessment results in a readily understandable format. [Kean-15]

Consider consequences of high stakes testing independently of test format. [Kean-18]

Monitor the use of tests for appropriateness. [Kean-18]

Take care that the use of alternative assessments do not create test bias problems. [Kean-18]

Special Populations

Special attention must be given to assessment results for homeless, migrant and N or D children. [CCSSO-18]

Use of available date from tests in child's native language is necessary in some circumstances. [CCSSO-18]

Wait until the end of 2nd grade before administering tests. [RHS:LA-2]

Omit 2nd and 3rd grades from standardized testing. [SPC-52]

Performance Measures

Careful development of scoring criteria and intensive training of judges are the key to establishing consistency of judgment. [OTA-242]

When only a few tasks are used there is a much higher risk that a child's score will be associated with the particular tasks and not generalize to the whole subject area that the test is meant to cover. This can be mitigated by sampling students and tasks if scores are not required for every student on every task. [OTA-242,243]

Benefits in curriculum and staff development may offset the higher costs associated with performance assessment. [OTA-245]

Data on the impacts of performance assessment on various subgroups is needed in considering whether to employ such measures in high stakes situations such as school accountability or student selection. [OTA-247]

If performance assessment is given a larger role in testing programs, teachers will need to be involved in designing tasks, administering and scoring tests, and placing test results into context. [OTA-248]

Writing assessment is now workable for all three major testing functions. Other methods of performance assessment (e.g., portfolios, exhibitions, experiments, and oral interviews) still represent relatively uncharted areas. They have potential for classroom instructional guidance and system monitoring through sampling. But much research is needed before they can be used for high-stakes applications in students selection and placement. [OTA-249].

Because performance assessment is at a developmental stage, encouraging states and districts to pool experience and resources is an appropriate policy goal. [OTA-249]

Augment NRTs with performance assessment, doing so gradually and give teachers a chance to buy into it. [RHS:LA-2]

New Technologies for Assessment

The Federal Government could continue to support research and development of a wide range of new models for testing in various ways, including earmarking resources in programs like Chapter 1 for research into how advanced technologies can improve testing. [OTA-279]

Use emerging technology to support periodic monitoring of student progress and use this information to improve instruction. [Kean-15]

SUMMARY RECOMMENDATIONS FOR CHAPTER 1 ASSESSMENT

1. Determine what Congress, national, state and district Chapter 1 administrators, teachers, parents, and communities need to know about student performance and knowledge.

2. Identify the appropriate types of assessment (e.g., tests) that are designed to provide the different types of information these groups need.

3. Administer these appropriate assessments to the minimum number of students, classrooms, grades, schools, or districts that are needed to provide reliable information. For example, USDE may not need information about students at every grade level, but districts might need information for every grade in which Chapter 1 services are provided.

4. Sampling should be used wherever appropriate and technically feasible.

5. Reduce the reliance of Chapter 1 assessment on current norm-referenced, standardized, multiple choice, achievement tests.

6. Performance-based measures should be used where feasible and appropriate since they provide a more valid assessment of student outcomes associated with curricula based on recent research and theoretical developments in cognitive psychology.

7. Assessment to meet accountability needs should be conducted at a national level by the Federal Government.

8. Assessment to determine where and how to improve Chapter 1 services should be conducted at the state and local level by SEAs and LEAs.

9. National accountability assessment should be linked to the national goals and curriculum standards.

10. State and local assessments of the effectiveness of Chapter 1 services for the purpose of improving these services should be based on state and local desired outcomes and curriculum standards.

11. Some states and school districts will need assistance in setting reasonably high goals or outcomes, constructing appropriate standards, and selecting or developing appropriate measures.

12. Assessment used to identify and select students for Chapter 1 services should be based on multiple measures and linked to the state and local desired outcomes and standards.

13. Assessment to diagnose individual student needs and monitor student progress in the classroom should be closely linked to the instruction provided by the teacher.

14. Use different types of assessment only for the purposes for which they have been shown to be reliable (consistent), valid (accurate), and unbiased (fair).

15. Present assessment results in a format that is easily understood by the audience for which they are intended.

16. Assessment results used to make high stakes decisions should be based on multiple measures.

17. Assessment methods should be designed to accommodate, rather than exclude, special populations, e.g., limited English proficient, homeless, early grade levels, migrant, and neglected or delinquent.

18. Developers of performance-based assessments should involve teachers in designing tasks, administering and scoring performance measures, and interpreting results for intended audiences.

19. Provide support for the examination and development of new technologies for assessing student performance and knowledge.